

Cherkasky – Fusionsproteine enthaltend Antikörperbinde-, Antigenbinde- und Mikrotubulibinde und Immunantwortauslösende Regionen

Die Erfindung betrifft die Bereiche der Tumorphysiologie und der Biotechnologie.

In der Tumortherapie stellen Operationen, Bestrahlung und Chemotherapie nach wie vor die entscheidenden Maßnahmen zur Therapie der Erkrankung dar. Bei der chemischen Tumortherapie (Chemotherapie) werden je nach Tumortyp meist Zytostatika unterschiedlicher Wirkungsart verwendet, so etwa Alkylantien, Nitrosoharnstoffverbindungen, Folsäureantagonisten, Pyrimidin- und Purinanaloga wie Fluorouracil, Antibiotika mit Wirkung auf die DNA-abhängige RNA-Polymerase oder Enzyme wie L-Asparaginase. Eine Gruppe von Cytostatika für die Chemotherapie sind die Mitosehemmstoffe wie etwa Taxol und Vinca-Alkaloide.

Auf Grund ihrer sehr guten Antitumor-Aktivität haben besonders die Mitosehemmstoffe in letzter Zeit verstärkte Beachtung gefunden. Die Mitosehemmer beeinflussen den Aufbau oder Abbau der aus Mikrotubuli bestehenden Teilungsspindele – und greifen somit an der Zellteilung an. Das bekannte Colchicin oder Vinca-Alkaloide binden an spezifischen Bindestellen des α - oder β -Tubulins – als Baustein der Mikrotubuli – und bewirken z.B. eine Hemmung des Aufbaus der Mikrotubuli. Andere Mitiosegifte – beispielsweise das Taxol – bewirkt deren Destabilisierung.

Mitose- oder Spindelgifte sind hochgradig toxisch und sind daher für therapeutisches Zwecke problematisch. Die Toxizität von Colchicin ist sogar so hoch, daß diese Substanz bislang gar nicht therapeutisch verwendet wird. Das aus Eiben (*Taxus*) isolierte Alkaloid Taxol ist derzeit Gegenstand intensiver Forschung.

Die meisten Mitosehemmer binden an das β -Tubulin der Mikrotubuli. Dazu weisen sie Bindungsstellen auf, deren unterschiedliche hohe Spezifität für eine Klassifizierung der Mitosehemmer herangezogen wird. So werden verschiedene Gruppen wie der Colchizin-Typ, der Taxan-Typ, der *Vinca* Alkaloid Typ oder der Rhyoxin Typ unterschieden.

Auf Grund der hohen Toxizität der Zytostatika ist eine Therapie mit diesen Substanzen mit vielen Nebenwirkungen verbunden, die für die betroffenen Patienten oft kaum erträglich sind. Daher wird seit vielen Jahren an der Verbesserung der Therapien mit der Zielsetzung der Vermeidung oder Reduzierung der Nebenwirkungen gearbeitet. Ein Ansatz dazu stellt der Versuch dar, die Wirkstoffe gezielt nur zu den zu therapierenden Zellen – d.h. zu den Zielzellen – zu lenken.

Eine Möglichkeit zur Verwirklichung dieses Ansatzes basiert im wesentlichen darauf, zelltypspezifische Epitope zu identifizieren, einen Epitop-spezifischen monoklonalen Antikörper zu erzeugen und den derart gewonnenen Antikörper oder Antigen-bindende Fragmente davon mit einem therapeutisch wirksamen Molekül zu koppeln. Ein derartiger Ansatz ist Gegenstand eines Forschungsprojekts der Universität von Kalifornien mit dem Ziel einer spezifischen Therapie von Brustkrebs (Sherie L. Morrison, Ph.D.: "Antibody Fusion Proteins for the

"Therapy of Breast Cancer", University of California, Los Angeles, 1997-1999). Hierbei wurden Antikörper gegen die brustkrebspezifischen Moleküle her2/neu und CEA verwendet und mit immunstimulierenden Molekülen verbunden, welche die Aktivität der T-Zellen stimulieren.

Obwohl dieser Ansatz mit dem Vorteil einer hohen therapeutischen Selektivität einhergeht, ist er in der Praxis nur unter großen Anstrengungen bei hohem Aufwand und langer Entwicklungsdauer umzusetzen, da zahlreiche Entwicklungsschritte zu seiner Realisierung erforderlich sind. Hierzu müssen zunächst für den jeweiligen Zelltyp spezifische Antigene isoliert werden. Da es sich bei diesen in der Regel um Proteinantigene handelt, werden im folgenden zellspezifische Epitope des Antigens ermittelt, die möglichst geringe Ähnlichkeiten zu Epitopen der Proteine anderer Zelltypen aufweisen. Dies ist erforderlich zur Vermeidung von Kreuzreaktivitäten der therapeutisch eingesetzten Antikörper. Anschließend erfolgt die Herstellung monoklonaler, gegen das jeweilige Antigen gerichteter Antikörper, die im weiteren aufwendigen Selektions- und/oder Mutageneseverfahren wie etwa Phage Display unterzogen werden müssen, um zu einem Antikörper möglichst hoher Spezifität, bzw. möglichst geringer Kreuzreaktivität zu gelangen.

Darüber hinaus ergeben sich häufig Schwierigkeiten bei der Herstellung des gebrauchsfertigen Therapeutikums, da ein nicht-humaner Antikörper modifiziert werden muß, um ohne hohes allergenes Potential eingesetzt werden zu können. Dazu können die variablen Regionen, insbesondere jedoch die *Complementary determining regions* (CDR) in ein humanes Antikörpergerüst eingesetzt, wobei im fertigen Therapeutikum unterschiedlich große antigenspezifische Elemente des therapeutischen Antikörpers, so etwa die antigenbindenden Fragmente (Fab) zum Einsatz kommen. Dabei handelt es sich in aller Regel um antigenspezifische Elemente, die mindestens aus zwei separaten Polypeptidketten bestehen. Die Herstellung dieser komplexen antigenspezifischen Elemente und ihre Verknüpfung mit dem eigentlich therapeutischen Molekül ist in der Praxis oft aufwendig und erfordert komplexe Expressionskonstrukte und entsprechend geeignete Wirtszellen.

Bekannt sind wissenschaftliche Arbeiten, in welchen Fusionsproteinen bestehend aus Liganden und Mikrotubuli - Bindedomänen, sowie aus Liganden, Mikrotubuli - Bindedomänen und Membranpenetrationsdomänen beschrieben sind. (DE 199 25 052.9; DE 101 61 899.9; DE 101 61 738.0; DE 101 61 739.9 und DE 101 62 870.6)

Die Nachteile bestehen darin dass erstens, keine zusätzliche Verstärkung der Wirkung durch Auslösung einer Immunantwort erfolgt und zweitens die Auswahl an Tumorspezifischen Antigenen und Antigenbinderegionen relativ gering ist und man benötigt eine Komplexierung mit Antikörper, um jeden beliebigen Antikörper modifizieren zu können, um ihr fähig zu machen direkt nach der Penetration oder Internalisierung Cytoskelett – Bestandteile bzw. Mikrotubuli zu binden bzw. zu fesseln.

Die Aufgabe der Erfindung besteht darin, effektive und selektive neuartige Fusionsproteinen und Fusionsprotein – Antikörper – Komplexe gegen unterschiedliche Arten der Leukämien und solide Tumoren zu entwickeln.

Die Aufgabe der Erfindung wird durch Fusionsproteinen, Cherkasky – Fusionsproteine genannt, enthaltend Antikörperbinde -, Antigenbinde -, Mikrotubulibinde – und Immunantwortauslösende Regionen gelöst. Die Selektivität wird durch Zell – oder Tumor spezifischen Liganden der Fusionsproteinen enthaltend zusätzlich Mikrotubuli – Binde und Immunantwort auslösenden Regionen sowie durch Antikörper der Fusionsprotein – Antikörper – Komplexe erreicht.

Die Fusionsproteinen, die mit Antikörpern Komplexe bilden enthalten Antikörper – Binde – und Mikrotubulibinderegionen.

Die Wirkung dieser Fusionsprotein – Antikörper – Komplexen besteht darin, Mikrotubuli – bzw. Zytoskelett zu finden bzw. zu fesseln nach dem die Tumorzellen durch hochaffine Antikörper ihre Zielzellen binden und penetrieren.

Die Antikörper Binderegion ist z.B. Staphylokokken protein A (SPA), extrazelluläre Region des Fc Rezeptors CD 64 etc.

Die Mikrotubuli - Binderegion ist z.B. Gephyrin, putativen Mikrotubulibindeprotein, FLJ 31424 Fis, MID - 1, MAP, Tau etc.

Diese Fusionsproteine können außerdem lange und superlange Spacer bzw. - Linkerregionen wie z.B. Polyglyzin oder Polyprolin enthalten, die mit Membranpenetrationsdomäne (MBD) oder Proteintransduktionsdomäne (PTD) fusioniert sind.

Die Fusionsproteine können auch Immunantwort - auslösenden Regionen wie z.B. Fc, B 7.1 oder B 7.2 enthalten, um die Wirkung der Komplexe zu erhöhen.

Die Fusionsproteine können Nukleinsäure – wie z.B. RecA oder Polysaccharid - Bindedomänen wie z.B. die Cellulose binderegion des CiPA enthalten, bzw. mit denen fusioniert werden, um durch Verdichtung bzw. erhöhte Konzentration die Wirkung zu verstärken.

Die Fusionsproteine können mit GFP oder anderen fluoreszenten Proteinen fusioniert werden , um ihre Wirkung optisch zu verfolgen oder ihre Konzentration durch Intensität der Fluoreszenz zu messen.

Außerdem können diese Fusionsproteine Gelenkregionen, wie z.B. Fünf - Glyzinregionen und mindestens eine GST-, Histag oder eine andere Region zur Durchführung der Affinitätsreinigung enthalten.

Die Fusionsproteine enthaltend Antigenbinde -, Mikrotubulibinde – und Immunantwort auslösenden Regionen entfalten ebenfalls eine Doppelwirkung.

Die Wirkung dieser Fusionsproteine besteht in der Tumorzellspezifischen Internalisierung und Hemmung der Zellteilung durch Bindung der Mikrotubuli sowie in der Auslösung einer Tumorzellspezifischer Immunantwort. Die Mikrotubulibinderegion kann z.B. aus Gephyrin, putativen Mikrotubulibindeprotein, FLJ 31424 Fis, Tau, MID - 1 oder MAP 1 ausgewählt werden und die Immunantwort auslösende Region ist z. B. Fc - Region eins IgG - Antikörpers, B 7.1 oder B7.2 Regionen zur Auslösung einer T - Zell - Reaktion.

Die Antigenenbinderegionen können vorzugsweise aus folgenden Proteinen - EGF, FGF, CSF, MGF, IL - 15, Il - 2, etc. ausgewählt werden.

Diese Fusionsproteinen können entweder eine Proteintransduktions (PTD) oder keine PTD enthalten. PTD ist dann nicht nötig, wenn die Antigenbinderegion ein Ligand darstellt welches nach der Interaktion mit dem entsprechenden Rezeptor in die Zelle internalisiert wird. In anderen Fällen dient PTD dazu Fusionskonstrukt in die Zielzellen zu bringen, bzw. zu internalisieren. Die eine PTD bzw. MPD ist z.B. die von Gen – 3 – Protein des Bakteriophagen fd, gp 41 oder Tat Protein HIV – 1.

Die beschriebene Proteintransduktionsdomäne (PTD) ist eine elf - Aminosäure lange Region, die eine Region des HIV Tat Proteins darstellt.

Dem Forscher Dowdy und seinen Kollegen ist gelungen, 60 Proteine in der Größen Ordnung zwischen 15 kDa und 120 kDa zu fusionieren und nach folgender Denaturierung der Fusion

mit Harnstoff ins Zytosol zu transportieren. (Science (285, 1569 - 1572, 1999) und Nature biotechnology Vol. 17 S. 942, Oct. 1999).

Nach der Internalisierung des Fusionsproteins wirkt die Mikrotubuli - Bindedomäne wie z.B. Gephyrin, Tau, MAP oder MID - 1 im Zytosol. Sie bindet Mikrotubuli und fesselt somit das Zytoskelett. Das dynamische Gleichgewicht (Wilde et al. Nature cell biology 2001, March, vol. 3 und Carazo. Salas et al. Nature Cell biology 2001, March, vol.3.) der Mikrotubuli wird beeinträchtigt und die jeweilige Zelle kann sich nicht mehr teilen. Sobald sie sich nicht mehr teilt, stirbt sie.

Dadurch wird das Wachstum des Tumors, z.B. eines soliden Tumors gehemmt.

Die Fusionsproteine können zusätzlich GFP oder eine andere flureszente Region enthalten um die Wirkung optisch zu verfolgen und die Konzentration in einer Lösung durch Intensität der Fluoreszenz zu messen. Ausserdem können diese Fusionsproteine Gelenkregionen, vorzugsweise Fünf - Glyzin - Region, und mindestens ein GST -, His tag oder eine andere Region zur Durchführung der Affinitätsreinigung enthalten.

In der Seq. 1a ist die Aminosäuresequenz des Fusionsproteins SPA – 5G -- Gephyrin dargestellt.

In der Seq. 1b ist die Nukleinsäure kodierend für dieses Fusionsprotein dargestellt.

In der Seq. 2a ist die Aminosäuresequenz des Fusionsproteins SPA – 5G – Mikrotubuli Bindepotein (MBP) dargestellt.

In der Seq. 2b ist die Nukleinsäuresequenz kodierend für das Fusionsprotein in der Fig. 2a dargestellt.

In der Seq. 3 ist die Nukleinsäure kodierend für das Fusionsprotein SPA – 5G – FLJ 31424 fis dargestellt.

In der Seq. 4 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 15 – 5G – Gephyrin - Fc dargestellt.

In der Seq. 5 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 2 – 5G – Gephyrin - Fc dargestellt.

In der Seq. 6 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 15 – 5G – MBP - Fc dargestellt.

In der Seq. 7 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 2 – 5G – MBP - Fc dargestellt.

In der Seq. 8 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 15 – 5G – FLJ 31424 fis - Fc dargestellt.

In der Seq. 9 ist die Nukleinsäuresequenz kodierend für das Fusionsprotein IL 2 – 5G – FLJ 314424 fis - Fc dargestellt.

In der Seq. 10 ist die Nukleinsäuresequenz des Fusionsproteins SPA – 5G – MBP - Fc dargestellt.

In der Seq. 11 ist die Nukleinsäuresequenz des Fusionsproteins SPA – 5G – Gephyrin - Fc dargestellt.

In der Seq. 12 ist die Nukleinsäuresequenz des Fusionsproteins SPA – 5G – FLJ 314424 fis - Fc dargestellt.

Die Wirkung der Fusionsproteine in den Sequenzen 1a – 3 besteht darin, dass sie einen beliebigen therapeutisch wirksamen Antikörper binden und dadurch modifizieren können. Die durch Bindung der Fusionsproteinen an Antikörper entstandenen Fusionsprotein – Antikörper – Komplexe können nach der Bindung der entsprechenden Antigenen und nach nachfolgender Internalisierung oder Penetration in den Zellen Mikrotubuli binden somit fesseln und dadurch die Zellteilung der bosartig veränderten Zellen hemmen.

Die Wirkung der Fusionsproteine in den Sequenzen 4 – 9 , besteht in der zielgerichteten Bindung ihrer Antigene oder Zell – bzw. Tumor spezifischen Rezeptoren, wonach die Doppelwirkung der Mikrotubulibinderegionen und zusätzlich die Induktion einer Immunreaktion durch Makrophagen entfalten wird bzw. erfolgt.

Die Wirkung der Fusionsproteine in den Sequenzen 10 – 12 , besteht in der Modifikation der Antikörper: die Komplexe besitzen auch eine Doppelwirkung und zwar die Fähigkeit Mikrotubuli zu binden und zusätzlich die Immunreaktion durch Aktivierung der Makrophagen zu induzieren.

Die DNA kodierend für die Fc Region des humanen Immunglobulins G ist von Nakamura, S., Sakugi, I., Kitai, K und Ichikawa, Y (NCBI <http://www.ncbi.nlm.nih.gov/entrez...> NCBI Sequence Viewer,) beschrieben.

Die humane mRNA für PMBP (putative microtubuli – binding protein) ist von Nadezhina E.S., beschrieben und ebenfalls in NCBI Sequence Viewer zu finden.

Die mRNA kodierend für Homosapiens Gephyrin (GPH) ist von NCBI beschrieben.

Die mRNA für Homo sapiens Interleukin – 2 ist von Chikara, S.K. und Sharma G beschrieben und ist in NCBI Sequence viewer zu finden.

Die Homo Sapiens mRNA für Interleukin 15 ist von Sorel, M.A. und Jacques, Y beschrieben wird in menschlichen Keratinocyten exprimiert und ist in NCBI Sequence Viewer zu finden.

Die Homo Sapiens cDNA für FLJ31421 fis, welche ein Mikrotubulibindeprotein kodiert, ist von Ota et al und Tashiro et al beschrieben und ist in NCBI Sequence Viewer zu finden.

Die SPA – Gensequenz bzw. die DNA kodierend für Staphylokokkenprotein A ist von El – Sayed, A., Alber J., Laemmer et al beschrieben und ist in NCBI Sequence Viewer zu finden.

Die Nukleinsäuren sind erhältlich von den Autoren. Alternativ dazu können die Sequenzen mit Hilfe von PCR und RT – PCR mit entsprechenden Primern herausamplifiziert und kloniert werden.

Diese Techniken sind dem Fachman bekannt.

Die Klonierung und Expression der Konstrukte wird vorzugsweise in E.Coli durchgeführt. Die frische E. Coli Zellkultur wird unter Zugabe von 75mM CaCl₂ (steril, kalt 250 ml) und Glycerin (seril, kalt 5mM) vorbereitet. Bei der Durchführung wird das Medium zuerst auf 37°C erwärmt, mit 8ml frischer E.Coli – Kultur angeimpft und bei 37°C kräftig geschüttelt. Danach wird die Kultur abgekühl. Die Zellen werden 10 Minuten lang bei 6000 rpm im GSA – Rotor der Sorvall – Zentrifuge abzentrifugiert und in 200 ml eiskaltem CaCl₂ 875 mM)

suspendiert. Danach werden die Zellen 20min auf Eis gestellt und wieder 10 min bei 6000 rpm in GSA – Rotor (Sorwall – Zentrifuge) abzentrifugiert.

Die Zellen werden in 20 ml eiskaltem CaCl₂ (75 mM) resuspendiert und wieder für 15 min auf Eis gestellt. Dazu wird 4,2 ml Glycerin zugegeben und gemischt. Die Lösung wird in sterile Eppendorfgefäße zu 0,5 ml abgefüllt und bei – 70°C eingefroren.

Die Transformation von E.Coli erfolgt mit 20 Mikroliter Ligationsansatz (oder max. 0,5 Mikroliter einer Plasmidpräparation, die 10 fach verdünnt wird). 100 ml kompetenter E.Coli Zellen werden zugegeben.

Danach werden diese ca. 30 min auf Eis (bis ca 1h) aufbewahrt, danach 2 – 3 min. auf 42 – 43°C erwärmt, um den Hitzeschock hervorzurufen und dann wieder auf Eis gestellt. Die Zellen werden z.B. auf X – Gal, Lbamp ausplattiert.

Die Zell – PCR erfolgt mit folgendem PCR – Mix pro Ansatz: 9 Mikroliter H₂O, 10ml 10xPCR – Puffer 10 Mikroliter dNTP – Mix (2 bis 2,5 mM jedes), 5ml BSA (20 mg/ml) oder 5 Mikroliter H₂O zusätzlich, 2ml 5' – Primer (20 pmol / Mikroliter); 2ml 3, - Primer (20 pmol / Mikroliter) und 2 ml Taq – Pol (54 / Mikroliter); insgesamt 40 Mikroliter.

Eine kleine Menge Zellen wird von der Platte abgenommen, und in einem 1,5 ml Eppendorf – Gefäß unten an der Wand verteilt.

Danach werden die Proben 2 min in 600 W Mikrowelle mit offenen Eppendorf – Gefäßen hingelegt.

Dann wird 200 Mikroliter H₂O zugegeben und gut gevortext, um zu resuspendieren. Danach wird 1 min in Eppendorf – Zentrifuge zentrifugiert. Vom Überstand werden 60 Mikroliter in ein PCR – Eppendorfgefäß gegeben und 40 Mikroliter PCR mix zugegeben.

Die PCR wird gestartet. Die Zyklen sind wie üblich, je nach annealing – Temperatur der Primer und Länge des zu Synthesieren den DNA – Fragmentes, z. B.

94°C	2 min	45°C	1 min	72°C	2 min	1x
94°C	30 sec	45°C	30sec	72°C	2 min	4x
94°C	30 sec	58°C	30sec	72°C	2 min	32x
94°C	30 sec	58°C	30sec	72°C	5 min	1x

Die Ligation wird mit 20 Mikroliter - Ansätzen gemacht : 15 Mikroliter zu ligierende DNA in H₂O, 4 Mikroliter 5x Ligase – Puffer mit PEG 1 Mikroliter T4 DNA – Ligase (1U/ Mikroliter), insgesamt 20 Mikroliter. Sticky – end – Ligierungen 1 – 2,5 h bei Raumtemperatur und blunt – end – ligierungen 4h bei Raumtemperatur. Ligase – Puffer besteht aus 250 mM tris/HCL pH 7, 6, 50 mM MgCl₂, 25% PEG 6000 (Sigma oder Serva) oder PEG 8000, 5mM ATP und 5mM DTT.

Die E. coli wird also transformiert.

Ausführungs - Beispiel 1

Klonierung und Expression des Fusionskonstruktes
N - IL - 15 - L - Gephyrin - Fc - C

c DNA für Gephyrin wird durch PCR kloniert, bzw. die Homosapiens Gephyrin (GPH) m RNA wird durch RT - PCR kloniert. Die Daten der GPH m RNA - Sequenz sind beim National Center für Biotechnology Information, NIH, Bethesda MD, 20 894, USA erhältlich, sowie auf der Internet - Seite von NCBI (<http://www.ncbi.nlm.nih.gov>) zu finden.

Die IL - 15 mRNA sowie die Fc vom IgG - mRNA werden ebenfalls durch RT - PCR kloniert.

Das Fusionsprotein bzw. das Fusionsprodukt wird aus PCR - Produkten zusammen gesetzt. Die PCR - Primern sind so konstruiert, dass sie Restriktionsstellen auf 5' und 3' Enden enthalten, um spätere Ligationsschritte durchzuführen. Die 5' und 3' Endes des IL - 15 PCR

Produkts enthalten Bam HI und Hind III Restriktionsstellen. Die 5' und 3' Enden des Gephyrin - PCR - Produktes enthalten EcoR I und KpnI Restriktionsstellen und die 5' und 3' Endes des Fc - PCR - Produktes enthalten Pst I und Sac I Restriktionsstellen.

Ligation der IL - 15, Gephyrin und Fc - sequenzen in den pUC 19 (2686 bp) - Vektor erfolgt unter Standard - bedingungen. Der PUC 19 - Vektor wird zuerst mit Bam HI und Hind 3 behandelt.

Der IL - 15 Segment wird durch diese Behandlung in den Vektor hineinligiert, pUC 19 - IL - 15 wird mit Eco RI und Kpn I behandelt, um den Gephyrin - Segment hineinzuligieren. Der pUC 19 - IL - 15 - Gephyrin - Vektor wird abschliessend mit Pst I und Sac I behandelt, um Fc Segment in den Vektor hineinzuligieren.

Ligationspuffer wird aus 66 mM Tris, pH 7,6,5 mM 5mM DTT und 1mM ATP sowie aus 20 Mikroliter T4 - DNA Ligase zusammengesetzt.

Das Ligationprodukt wird in E. Coli transformiert, exprimiert und abschliessend gereinigt.

Ausführungs - Beispiel 2

Klonierung und Expression des Fusionskonstruktes SPA - 5G - Gephyrin

C DNA für Gephyrin wird durch PCR kloniert, bzw. die Homo sapiens Gephyrin (GPH) m RNA wird durch RT - PCR kloniert. Die Daten der GPH m RNA - Sequenz sind bei im Internet, beim National Center for Biotechnology information, NIH, Bethesda MD 208 94, USA erhältlich, sowie auf der Internet Seite von NCBI (<http://www.ncbi.nlm.nih.gov:80/entrez/eotide...>) zu finden.

C DNA für SPA ligiert mit dem Primer kodierend für die Fünf - Glyzin - Spacer wird ebenfalls mit PCR kloniert.

Das Fusionsprotein wird aus PCR - Produkten zusammengesetzt. Die PCR - Primern sind so konstruiert, dass sie Restriktionsstellen auf 5' und 3' Enden enthalten, um spätere Ligationsschritte durchzuführen. Die 5' und 3' Enden des Gephyrin - PCR - Produkts enthalten Bam HI und Hind III Restriktionsstellen. Die 5' und 3' Enden des SPA - PCR - Produkts enthalten XmnI und Bg III Restriktionsstellen.

Nach der Amplifikation und Reinigung werden die PCR Produkten in PCR II Vektoren ligiert. Positive Klone werden durch Screening Plasmide richtiger Masse identifiziert. Die Klone werden durch DNA - Sequenzierung bzw. durch Standardmethoden überprüft bzw. bestätigt.

Das Gephyrin - PCR - Produkt wird aus der PCR II durch restriktive Spaltung durch Bam HI und Hind III herausgeschritten, und SPA - PCR - Produkt wird aus dem PCR II durch Xmn I und Bg III herausgeschnitten.

Ligation der Gephyrin und SPA - Segmente in den pMal - c 2 Expressionsvektor erfolgt unter Standard - Bedingungen. Der p Mal - c 2 Vektor wird mit Bam H I und Hind 3 behandelt. Der Gephyrin - Segment wird durch diese Behandlung in den pMal - c 2 hinein ligiert.

PMal - c 2 - Gephyrin wird mit X mnI und Bam HI geschnitten, um SPA Segment hinein zu ligieren.

Ligationspuffer wird aus 66 mM Tris PH 7,6, 5 mM Mg Cl₂, 5 mM DTT und 1 mM ATP, sowie aus der T4 DNA Ligase (insgesamt 20 Mikroliter) zusammengesetzt. Die Ligation wird bei 14°C durchgeführt.

Das Ligationsprodukt wird in E Coli transformiert, exprimiert und abschliessend gereinigt.

Mit den in den Ausführungsbeispielen 1 und 2 beschriebenen Methoden sowie mit den anderen dem Fachman bekannten Methoden werden die Sequenzen in den Figuren 1 – 12 kloniert und exprimiert.

Cherkasky – Fusionsproteine enthaltend Antikörperbinde-, Antigenbinde- und Mikrotubulibinde und Immunantwortauslösende Regionen

Patentansprüche

Beansprucht werden:

1. Fusionsproteine, dadurch gekennzeichnet, dass sie spezifische Antigenbinderegionen wie vorzugsweise EGF, FGF, CSF, MGF, IL – 15, IL – 2, oder anderen Liganden oder ihre Regionen, Mikrotubulibinderegionen wie vorzugsweise Gephyrin, Tau, MAP, MID – 1, MBP, bzw. put MBP oder PMBP, FLJ 31424 Fis, oder ihre Regionen und Immunantwortauslösenden Regionen wie vorzugsweise Fc des IgG, B 7.1, B 7.2 oder ihre Regionen, enthalten.
2. Fusionsproteine, dadurch gekennzeichnet, dass sie Antikörperbinderegionen, wie vorzugsweise Staphylokokkenprotein A (SPA), extra zelluläre Region des Fc Rezeptors CD 64 oder ihre Regionen und Mikrotubulibinderegionen, vorzugsweise Gephyrin, Tau, MAP, MID – 1 MBP, FLJ 31424 Fis, oder ihre Regionen enthalten.
3. Fusionsproteine nach dem Anspruch 2 gekennzeichnet durch immunauslösende Regionen wie vorzugsweise Fc Regionen des Immunglobulins G, oder IgG, HLA – B 7.1 oder HLA – B 7.2.
4. Fusionsproteine nach den Ansprüchen 1 – 3, gekennzeichnet durch lange und sehr lange Spacer oder Linker Regionen wie vorzugsweise Polyglycin, Polyprolin oder Spacer die Glycine und Proline enthalten.
5. Fusionsproteine nach den Ansprüchen 1 – 4, gekennzeichnet durch mindestens eine Nukleinsäurebinderegion vorzugsweise RecA.
6. Fusionsproteine nach den Ansprüchen 1 – 5, gekennzeichnet durch mindestens eine Polysaccharidbinderegion wie vorzugsweise die Cellulosebinderegion des CipA Proteins.
7. Fusionsproteine nach den Ansprüchen 1 – 6, dadurch gekennzeichnet dass sie GFP, eine andere fluoreszente Region, Membranpenetrationsdomäne wie vorzugsweise das Gen – 3 – Protein des Bakteriophagen fd, gp 41 oder Tat Protein des HIV – 1 oder eine mindestens andere Region, enthalten bzw. umfassen.
8. Fusionsproteine nach den Ansprüchen 1 – 7, gekennzeichnet durch eine GST Region, His tag oder eine andere Region zur Durchführung der Affinitätsreinigung.

9. Fusionsproteine nach den Ansprüchen 1 – 8, gekennzeichnet durch eine Spaltstelle für eine Protease vorzugsweise zum Abschneiden der GST oder eines anderen Reinigungstags vom Konstrukt.
10. Nukleinsäure und Aminosäure Sequenzen, DNA – Vektoren, Klonierungs – und Expressionssysteme für die Fusionsproteine nach den Ansprüchen 1 – 9, sowie alle Aminosäure und Nukleinsäuresequenzen oder Sequenzen 1 – 12 .

Seq. 1a

AAQHDEAQQNAFYQVLNMPNLNADQRNGFIQSLKDDPSQSANVL
GEAKKLNESQAPKADNNFNKEQQNAFYEILNMPNLNEEQRNGFIQSLKDDPSQSANLL
SEAKKLNESQAPKADNKFNKEQQNAFYEILHLPNLNEEQRNGFIQSLKDDPSQSANLL
AEAKKLNDAAQAPKADNKFNKEQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEIL
AEAKKLNDAAQAPKEEDNNKPGKEDGNKPGKEDGNGGGGG
MSPFPLTSMDKAFITVLEMTPVLGTEI:NYRDGMGRVLAQDVYA
KDNLPPFPASVKDGYAVRAADGPGDRFIIGESQAGEQPTQTVMMPQVMRVTTGAPIPC
GADAVVQVEDTELIRESDDGTEELEVRILVQARPGQDIRPIGHDIKRGECLAKGTHM
GPSEIGLLATVGVTEVEVNKFVVAVMSTGNELLNPEDDLLPGKIRDSNRSTLLATIQ
EHGYPTINLGIVGDPDDLLNALNEGISRADIITSGGVSMGEKDYLKVLDIDLHAQ
IHFGRVFMKPGLPTTFATLDIDGVRKIIFALPGNPVSAVTCNLFVVPALRKMQGILD
PRPTIIKARLSCDVKLDPRPEYHRCILTWHHQEPLPWAQSTGNQMSSRLMSMRSANGL
LMLPPKTEQYVELHKGEVVDMVIGRL

Seq. 1b

tgctgcgcaa cacgatgaag ctcaacaaaa cgcttttat caagtctaa atatgcctaa
cttaaatgt gatcaacgca atggtttat ccaaagcctt aaagatgatc caagccaaag
tgctaacgtt ttaggtgaag ctaaaaattt aaacgaatct caagcaccga aagctgacaa
caattcaac aaagaacaac aaaatgctt ctatgaaatc ttgaacatgc ctaactgaa
cgaagaacaa cgcaatggtt tcattccaaag cttaaaagat gacccaagtc aaagtctaa
cctattgtca gaagctaaaa agttaaatga atctcaagca ccggaaagcgg ataacaatt
caacaaagaa caacaaaatg cttctatga aatcttacat ttacctaact taaacgaaga
acaacgcaat ggttcatcc aaagcctaaa agatgaccca agccaaagcg ctaacccttt
agcagaagct aaaaagctt atgatgcaca agcacaaaaa gctgacaaca aattcaacaa
agaacaacaa aatgcttct atgaaatttt acatttacct aacttaactg aagagcaacg
taacggcttc atccaaagcc taaaagacga tccttcagtg agcaaagaaa ttttagcaga
agctaaaaag ctaaacgtt ctcaaggacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc aacaaacctg gcaaagaaga cggtaacggc ggc ggc ggc
gtttaggtca cagtgtgtc gatatcacca aggtggctag aagacatcgc atgttcctt
ttcctctgac atctatggac aaagccctta tcacagtccct ggagatgact ccggtgcttg
ggacagaaaat catcaattac cgagatggaa tggggcgagt cttgctcaa gatgtatatg
caaaagacaa ttacccccc ttcccagcat cagtaaaaga tggctatgct gtccgagctg
ctgatggccc aggagatcgt ttcatcattt gggaaatccca agctggtaa cagccaaactc
agacagtaat gccaggacaa gtcacgcggg ttacaacagg tgctccaata ccctgcggtg
ctgatgcagt agtacaagtg gaagataccg aacttatcag ggaatcagat gatggcactg

aagaacttga agtgcgaatt ctggtgcaag ctcggccagg ccaagatatc agacccatcg
gccatgacat taaaagaggg gaatgtttt tggccaaagg aaccacatg ggcccctcag
agattggtct tctggcaact gtaggtgtca cagaggttga agttaataag ttccagtgg
ttcagtcgt gtcaacacagg aatgagctgc taaatcctga agatgaccc tcaccaggaa
agattcgaga cagcaatcgt tcaactcttc tagcaacaat tcaggaacat gtttacccca
cgatcaactt gggtatttta ggagacaacc cagatgactt actcaatgcc ttgaatgagg
gtatcagtcg tgctgtatgtc atcatcacat caggggggtgt atccatgggg gaaaaggact
atctcaagca ggtgctggac attgatcttc atgctcagat ccattttggc agggtttta
tggaaaccagg cttgccaaca acatttgc当地 cttggatat tgatgggttta agaaaaataa
tctttgcact acctggaaat cctgtatcg ctgtggtcac ctgcaatctc ttgttgtgc
ctgcactgag gaaaatgcag ggcatttgg atcctcggcc aac:catcatc aaagcaaggt
tatcatgtga tgtaaaactt gatcctcgtc cagaatacca tcgggtata ctaacttggc
atcaccaaga accactacct tggcacaga gtacaggtaa tcaaatgagc agccgtctga
tgagcatgct cagtgccaaat ggattgttga tgctaccccttcc aaagacagaa cagtacgtgg
agctccacaa aggcgaggtt gtggatgtca tggcattgg acggctatga tggcaccag

Seq. 2a

AAQHDEAQQQNAFYQLNMPNLNADQRNGFIQSLKDDPSQSANVL
GEAKKLNESQAPKADNNFNKEQQQNAFYEILNMPNLNEEQRNGFIQSLKDDPSQSANLL
SEAKKLNESQAPKADNKFNKEQQQNAFYEILHLPNLNEEQRNGFIQSLKDDPSQSANLL
AEAKKLNDDAQAPKADNKFNKEQQQNAFYEILHLPNLTEEQRNGFIQSLKDDPSVSKEIL
AEAKKLNDDAQAPKEEDNNKPGKEDGNKPGKEDGNGGGGG
AAASTAXASTAKETAEAVALDITAYP

Seq. 2b

tgctgcgcaa cacgatgaag ctcaacaaaa cgcttttat caagtcttaa atatgcctaa
cttaaatgct gatcaacgca atggtttat ccaaagcctt aaagatgatc caagccaaag
tgctaacgtt ttaggtgaag ctaaaaaatt aaacgaatct caagcacccga aagctgacaa
caattcaac aaagaacaac aaaatgcttt ctatgaaatc ttgaacatgc ctaacttgaa
cgaagaacaa cgcaatggtt tcatccaaag cttaaaagat gacccaaatc aaagtgctaa
cctattgtca gaagctaaaa agttaaatga atctcaagca ccgaaagcgg ataacaaatt
caacaaagaa caacaaaatg ctttctatga aatcttacat ttucttaact taaacgaaga
acaacgcaat gtttcatcc aaagcctaaa agatgacccca agccaaagcgt ctaaccttt
agcagaagct aaaaagctaa atgatgcaca agcaccaaaa gctgacaaca aattcaacaa
agaacaacaa aatgcttct atgaaatttt acattacct aacttaactg aagagcaacg
acqqacttc atccaaagcc taaaqacga tccttcagtg agccaaagaaa tttaqcaga

agctaaaaag ctaaacgatg ctcaagcacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc aacaaacctg gcaaagaaga cgtaacggc ggc ggc ggc ggc
gcggccgcgt cgaccgcgn cgctcgacg gcaaaggaga ctgctgaggc ttttgctgat
ganatactgg anaaggctgg gccacttgtt gctgtgtctg ctgtgcact tgatataact
gcctaccctt aaaagccaaa

Seq. 3

tgctgcgcaa cacgtgaag ctcaacaaaa cgcttttat caagtctaa atatgcctaa
cttaaatgct gatcaacgca atggtttat ccaaagccit aaagatgatc caagccaaag
tgctaacgtt ttaggtgaag ctaaaaaatt aaacgaatct caagcaccga aagctgacaa
caattcaac aaagaacaac aaaatgctt ctatgaaatc ttgaacatgc ctaacttgaa
cgaagaacaa cgcaatggtt tcatccaaag cttaaaagat gacccaagtc aaagtctaa
cctattgtca gaagctaaaa agttaatga atctcaagca ccgaaagcgg ataacaaatt
caacaaagaa caacaaaatg cttctatga aatcttacat ttacctaact taaacgaaga
acaacgcaat ggttcatcc aaagcctaaa agatgaccca agccaaagcg ctaaccttt
agcagaagct aaaaagctaa atgatgcaca agcaccaaaa gctgacaaca aattcaacaa
agaacaacaa aatgcittct atgaaatttt acatttacct aacttaactg aagagcaacg
taacggcttc atccaaagcc ttaaagacga tccttcagtg agcaaagaaa ttttagcaga
agctaaaaag ctaaacgatg ctcaagcacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc aacaaacctg gcaaagaaga cgtaacggcggcggcggc
aatgtcccga attcccagcc tcaccacccc ttctcagtaa tgacccttgt tggttgcagg
aggtacctac tccatactga gggtaattt aagggaaaggc aaagtccagg cacaagagt
ggaccccagc ctctcactct cagttccact catccaactg ggaccctcac cacgaatctc
atgatctgat tcggltccct gtctccctt cccgtcacag atgtgagcca gggcactgct
cagctgtac cctaggtgtt tctgccttgt tgacatggag agagccctt cccctgagaa
ggcctggccc ctcctgtgc tgagccaca gcagcaggct gggtgtctg gttgtcagtg
gtggcaccag gatgaaaggc caaggcaccc agggcaggcc cacagtcccg ctgtccccca
ctgcacccct agctttagc tgccaaacctc ccagacagcc cagccgcgt ctcagctcca
catgcataatc atcagccctc cacacccgac aaaggaaac acacccctt ggaaatggtt
ttttcccccc agtcccagct ggaagccatg ctgtctgttc tgctggagca gctgaacata
tacatagatg ttgccctgcc ctcccatct gcaccctgtt gagtttagt tggattgtc
tgtttatgt tggattcacc agagtacta tgatagtgaa aaaaaaaaaa aaaaaaaaaa
aggacgcacg tatctgaaa tgcttgaaa gaggttcta acccaccctc acgagggtgtc
tctcaccctt acactggac tcgtgtggcc tgtgtggc caccctgctg gggcctccca
agtttgaaa ggcttcctc agcacctggg acccaacaga gaccagctc tagcagctaa
ggaggccgtt cagctgtac gaaggcctga agcacaggat taggactgaa gcgatgtgt
cccctccctt actccctt gggctccct gtgtcaggc acgtactagg tcttggtgt
ggtctggctt gcggcgcgag gatggttctc tctggcata gcccgaagtc tcatggcagt
ccaaaggag gcttacaact ctcgtacac aaaaaagg aagccactgc cagctgggg

Seq. 4

ccaaaccagg ctacttgtt ttacttctaa acagtcattt tctaactgaa gctggcattc
atgtttcat tttgggctgt ttcagtgcag ggcttcctaa aacagaagcc aactgggtga
atgttaataag tgatttgaaa aaaattgaag atcttattca atctatgcat attgatgcta
ttatataac ggaaaagtgtat gttcacccccca gttgcaaagt aacagcaatg aagtgcattc
tttggagtt acaagttatt tcacttgagt ccggagatgc aagtattcat gatacaqtag

aaaatctgat catcctagca aacaacagt tgcattcaa tggaatgt a cagaatctg
gatgcaaaga atgtgaggaa cttagggaaa aaaatattaa agaattttg cagagtttgc
tacatattgt ccaaatgttc atcaacactt ctt ggc ggc ggc ggc
gtttaggtca cagtgcgtc gatatcacca aggtggctag aagacatgc atgttcctt
ttcctctgac atctatggac aaagccctta tcacagtcc ggagatgact ccggtgctt
ggacagaaaat catcaattac cgagatggaa tggggcgagt cctgctcaa gatgtatatg
caaaagacaa ttaccccccc ttccagcat cagaaaaaaga tggctatgt gtccgagctg
ctgatggccc aggagatgt ttcatcattt gggaaatccca agctggtaa cagccaactc
agacagtaat gccaggacaa gtcatgcggg ttacaacagg tgctccaata ccctgcggtg
ctgatgcagt agtacaatgt gaagataccg aacttatcag ggaatcagat gatggactg
aagaactga agtgcgaatt ctggtgcag ctcggccagg ccaagatatc agacccatcg
gccatgacat taaaagaggg gaatgtttt tggccaaagg aacccacatg ggcccctcag
agattggct tctggcaact gtaggtgtca cagaggtga agttaataag ttccagtg
ttgcagtcat gtcaacaggg aatgagctgc taaatcctga agatgaccc tcaccaggga
agattcgaga cagcaatcgt tcaactcttc tagcaacaat tcaggaacat gttacccca
cgatcaactt ggttattgtt ggagacaacc cagatgactt actcaatgcc ttgaatgagg
gtatcagtgc tgctgatgtc atcatcacat caggggggtt atccatgggg gaaaaggact
atctcaagca ggtgctggac attgatcttc atgctcagat ccaatttggc agggtttta
tgaaccagg ctgccaaca acatttgcaa cttggatat tggatgtt agaaaaataa
tcttgcaact acctggaaat cctgtatcgg ctgtggtcac ctgcaatctc ttgttgtgc
ctgcactgag gaaaatgcag ggcacatctgg atcctcgcc aaccatcatc aaagcaagg
tatcatgttga tgtaaaactt gatcctcgatc cagaataccca tcggtgtata ctaacttggc
atcaccaaga accactacctt tggcacaga gtacaggtaa tcaaattgagc agccgtctga
tgagcatgcg cagtgccaaat ggattgttga tgctacctcc aaagacagaa cagtgatgg
agctccacaa aggcgagggtt gtggatgtca tggcattgg acggctatga tggcaccag
ctgttgcacaa ttaatcatcg gctcgatataa tggatgttga tggatgttgc taacaattt
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctgg
gggaccgtca gtcttcctt tccccccaaa acccaaggac accctcatga tctccggac
ccctgagggtc acatgcgtgg tggatgttgc gagccacgaa gaccctgagg tcaagttca
ctggatgttgc gacggcgtgg aggtgcataa tgccaaagaca aagccgcggg aggaggat
caacagcacg taccgggtgg tcagcgttcc caccgtccgt caccaggact ggctgaatgg
caaggatgttgc aagtgcagg tctccaaacaa agccctccca gccccatcg agaaaaaccat
ctccaaagcc aaaggcgaccc cccgagaacc acaggtgtac accctgcccc catccggga
ggagatgacc aagaaccagg tcagcgttgc ctgcctggc aaaggcttct atcccgac
catcgccgtg gagtgggaga gcaatggca gcccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttcttcctt ctatagcaag ctcaccgtgg acaagagc
gtggcagcag gggaaacgttct tctcatgttc cgtgatgttgc gaggctgtc acaaccacta
cacgcagaag agcctctccc tggccggg taaataatag gatcc

Seq. 5

atgcctactt caagttctac aaagaaaaaca cagctacaac tggagcattt actgctggat
ttacagatga tttgaatgg aattaataat tacaagaatc ccaaactcac caggatgctc
acatttaagt tttacatgcc caagaaggcc acagaactga aacatctca gtgtcttagaa
gaagaactca aacctctgga ggaagtgcta aatttagctc aaagaaaaaa ctttcactta
agaccagg acttaatcag caatatcaac gtaatagttc tggactaaa gggatctgaa
acaacattca tgtgtgaata tgctgatgag acagcaacca ttgtagaatt tctgaacaga
tggattacct tttctcaaag catcatctca acactgactt gataa ggc ggc ggc ggc
gtttaggtca cagtgctgtc gatatcacca aggtggctag aagacatcgc atgtccctt
ttccctcgac atctatggac aaagcctta tcacagtcc ggagatgact ccggtgcttg
ggacagaaaat catcaattac cgagatggaa tggggcgagt cttgctcaa gatgtatatg
caaaagacaa ttaccccccc ttcccagcat cagtaaaaga tggctatgct gtccgagctg
ctgatggccc aggagatcgt ttcatcattt gggaatccca agctggtaa cagccaactc
agacagtaat gccaggacaa gtcatcgccc ttacaacagg tgctccaata ccctcggtg
ctgatgcagt agtacaagtg gaagataccg aacttatcag ggaatcagat gatggcactg
aagaacttga agtgcgaatt ctggtcaag ctcggccagg ccaagatatc agacccatcg
gccatgacat taaaagaggg gaatgtttt tggccaaagg aacccacatg ggccccctcag
agattggct tctggcaact gtaggtgtca cagaggtta agttaataag tttccagtgg
ttgcagtcat gtcaacaggg aatgagctgc taaatcctga agatgaccc ttaccaggga
agattcgaga cagcaatcgt tcaactcttc tagcaacaat tcaggaacat ggttacccca
cgatcaactt gggattgtt ggagacaacc cagatgactt actcaatgcc ttgaatgagg
gtatcagtgc tgctgatgtc atcatcacat cagggggtgt atccatgggg gaaaaggact
atctcaagca ggtgctggac attgatcttc atgctcagat ccatttggc agggtttttta
tggaaaccagg ctggccaaaca acattgcaa cttggatat tggatgttta agaaaaataaa
tcttgcaact acctggaaat cctgtatcgg ctgtggcac ctgcaatctc ttgttgtgc
ctgcactgag gaaaatgcag ggcacattgg atcctcgcc aaccatcatc aaagcaaggt
tatcatgtga tgtaaaactt gatcctcgac cagaatacca tcggtgtata ctaacttggc
atcaccaaga accactaccc tggcacaga gtacaggtaa tcaaattgagc agccgtctga
tgagcatgcg cagtgcacat ggattgttga tgctaccccttcc aaagacagaa cagtcgtgg
agctccacaa aggcgagggtg gtggatgtca tggcattgg acggctatga tggcaccag
ctgttgacaa ttaatcatcg gctcgatataa tggatgttgaat tggatgttgcgaa taacaatttc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactccctgg
gggaccgtca gtcttccttcc tccccccaaa acccaaggac accctcatga tctccggac
ccctgagggtc acatgcgtgg tggatgttgcgat gagccacgaa caccctgagg tcaagtttca
ctggtacgtg gacggcgtgg aggtgcataa tgccaaagaca aagccgcggg aggagcgt
caacagcacg taccgggtgg tcagcgtcc caccgtccgt caccaggact ggctgaatgg
caaggagtac aagtgcagg tctccaacaa agccctccca gccccatcg agaaaaaccat
ctccaaagcc aaaggcagc cccgagaacc acaggtgtac accctgcccc catccggga
ggagatgacc aagaaccagg tcagcctgac ctgcctggc aaaggcttct atccagcga

catgccgtg gagtgggaga gcaatggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ctttccttctt ctagcaag ctcaccgtgg acaagagcag
gtggcagcag gggAACgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 6

ccatccagtg ctacttgtt tacttctaa acagtcattt tctaactgaa gctggcattc
atgtcttcat ttggggctgt ttcagtgcag ggcttcctaa aacagaagcc aactgggtga
atgtaataag tgattgaaa aaaattgaag atcttattca atctatgcat attgatgcta
ctttatatac ggaaagtgtat gttcacccca gttgcaaagt aacagcaatg aagtgcattc
tctggagtt acaagttatt tcacttgagt ccggagatgc aagtattcat gatacagtag
aaaatctgat catcctagca aacaacagtt tgtcttctaa tggaaatgta acagaatctg
gatgcaaaga atgtgaggaa ctagaggaaa aaaatattaa agaattttg cagagtttg
tacatattgt ccaaattgttc atcaacactt ctt ggc ggc ggc ggc
gcggccgcgt cgaccgcgn cgcgtcgacg gcaaaggaga ctgctgaggc tggctgtat
ganatactgg anaaggctgg gccacttggt gctgtgtctg ctgtgcact tgatataact
gcctaccctt aaaagccaaa
ctgtgacaa ttaatcatcg gctcgatataa tgtgtggaat tgtgagcggtaaacaatttc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg
gggaccgtca gtcttcctct tccccccaaa acccaaggac accctcatga tctccggac
ccctgaggc acatgcgtgg tggtggacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtagtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagtac aagtgcagg tctccaacaa agccctccc gccccatcg agaaaaccat
ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc catccggga
ggagatgacc aagaaccagg tcagcctgac ctgcctggc aaaggcttct atcccagcga
catcgccgtg gagtgggaga gcaatggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttcttcct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggAACGTCT tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tggccggg taaataatag gatcc

Seq. 7

atgcctactt caagttctac aaagaaaaaca cagctacaac tggagcattt actgctggat
ttacagatga ttttgtaatgg aattaataat tacaagaatc ccaaactcac caggatgctc
acatttaagt ttacatgcc caagaaggcc acagaactga aacatttca gtgtcttagaa
gaagaactca aacctctgga ggaagtgcta aatttagctc aaagcaaaaaa ctttcactta
agacccaggg acttaatcag caatatcaac gtaatagttc tggaactaaa gggatctgaa
acaacattca tgtgtgaata tgctgatgag acagcaacca ttgtagaatt tctgaacaga
tggattacct tttctcaaag catcatctca acactgactt gataaggcggcggcggcggc

gcggccgcgt cgaccgcggn cgcgctgacg gcaaaggaga ctgctgaggc ttttgctgat
ganatactgg anaaggctgg gccacttgtt gctgtgtctg ctgttgact tgatataact
gcctaccctt aaaagccaaa
ctgtgacaa ttaatcatcg gctcgataa tgtgtggaaat tgtgagcgga taacaattc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctgg
gggaccgtca gtcttccttccccccaaa acccaaggac accctcatga tctccggac
ccctgaggc acatgcgtgg tggtgacgt gagccacgaa gaccctgagg tcaagtcaa
ctggtagctg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagttac aagtgcagg tctccaacaa agccctccc gccccatcg agaaaaaccat
ctccaaagcc aaagggcagc cccgagaacc acagggtac accctgcccc catccggga
ggagatgacc aagaaccagg tcagcctgac ctgcctggc aaaggcttct atcccagcga
catcgccgtg gagtgggaga gcaatggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ctttttcct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggAACGTC tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tttccccggg taaataatag gatcc

Seq. 8

ccatccagtg ctacttgcgt ttacttctaa acagtcattt tctaactgaa gctggcattc
atgtctcat ttgggctgt ttcaagtgcag ggcttcctaa aacagaagcc aactgggtga
atgtataag tgattgaaa aaaattgaag atcttattca atctatgcat attgatgcta
ctttatatac gggaaagtgtat gttcacccca gttcaaatg aacagcaatg aagtgcattc
tcttgagtt acaagttatt tcacttgagt ccggagatgc aagtattcat gatacagtag
aaaatctgtat catcctagca aacaacagtt tgcctctaa tggaaatgta acagaatctg
gatgcaaaaga atgtgaggaa ctagaggaaa aaaatattaa agaatttttgc cagagtttgc
tacatattgt ccaaatttttc atcaacactt ctt ggc ggc ggc ggc
aatgtcccga attcccaagcc tcaccaccc ttctcagtaa tgaccctgg tggttgcagg
aggtaacctac tccatactga gggtaaaatt aaggaaaggc aaagtccagg cacaagagtg
ggaccccagc ctctcactct cagttccact catccaactg ggaccctcac cacgaatctc
atgatctgtat tcggttccct gtcttcctt cccgtcacag atgtgagcca gggcactgct
cagctgtac cctagggttt tctgcctgt tgacatggag agagccctt cccctgagaa
ggcctggccc ttccctgtgc tgagccaca gcagcaggct ggggtcttg gttgtcagt
gtggcaccag gatgaaaggc caaggcaccc agggcaggcc cacagtcccg ctgtccccca
ttgcaccctt agcttgtac tgccaaacctc ccagacagcc cagcccgctg ctcagctcca
catgcatagt atcagccctc cacacccgac aaaggaaac acacccctt ggaaatggtt
ctttccccc agtcccagct ggaagccatg ctgtctgttc tgctggagca gctgaacata
tacatagatg ttgcctgcc ctccccatct gcaccctgtt gagttgtatg tggatttgc
tgtttatgtatg tggattcacc agagtacta tgatagtgaa aaaaaaaaaaaaaaaa

aggacgcatg tatcttgaaa tgcttgtaaa gaggttcta acccaccctc acgagggtgc
tctcacccccc acactggac tcgtgtggcc tgtgtgggtgc caccctgctg gggcctccca
agtttgaaa ggcttcctc agcacctggg acccaacaga gaccagcttc tagcagctaa
ggaggccgtt cagctgtgac gaaggcctga agcacaggat taggactgaa gcgtatgt
ccccctccct acttcccctt ggggctccct gtgtcagggc acagactagg tcttgtggct
ggtctggctt gcggcgcgag gatggttctc tctggtcata gccccaagtc tcatggcagt
cccaaaggag gcttacaact cctgcacac aagaaaaagg aagccactgc cagctgggg
gatctgcagc tcccagaagc tccgtgagcc tcagccaccc ctcagactgg gttcctctcc
aagctcgccc tctggagggg cagcgcagcc tcccaccaag ggccctgcga ccacagcagg
gattggatg aattgcctgt cctggatctg ctctagaggc ccaagctgcc tgctgagga
aggatgactt gacaagtcag gagacactgt tcccaaagcc ttgaccagag cacctcagcc
cgctgacctt gcacaaaactc catctgctgc catgagaaaa gggaaagccgc cttgcaaaa
cattgctgcc taaagaaaact cagcagcctc aggcccaatt ctgccacttc tggttggt
acagttaaag gcaaccctga gggacttggc agtagaaatc cagggcctcc cctggggctg
gcagcttcgt gtgcagctag agcttacctt gaaaggaagt ctctgggccc agaactctcc
accaagagcc tccctgccgt tcgctgagtc ccagcaattc tcctaagttg aaggatctg
agaaggagaa ggaaatgtgg ggttagattt gttgtggta gagatatgcc cccctcatta
ctgccaacag ttcggctgc atttcttcac gcacccctcggt tcccttcctt gaagttctt
tgccctgctc ttccacca tggccttct tatacggaaag gctctggat ctccccctt
tggggcaggg tcttggggcc agcctaagat catggtttag ggtgatcagt gctggcagat
aaattgaaaa ggcacgctgg ctgtgtatct taaatgagga caatcccccc agggctggc
actccctcccc tcccctcaact tctcccacct gcagagccag tgtcctggg tggcttagat
aggatataact gtatgccggc tcctcaagc tgctgactca ctttatcaat agttccattt
aaattgactt cagtggtagactgtatccct gttgttatt gctgttgta ctatgggg
aggggggagg aatgtgtaaat atagttaaca tggccaaagg gagatctgg ggtgcagcac
ttaaaactgcc tcgttaaccct tttcatgatt tcaaccacat ttgcttagagg gagggagcag
ccacggagtt agaggccctt ggggttctc tttccactg acaggcttc ccaggcagct
ggctagttca ttccctcccc agccaggtgc aggcgttagga atatggacat ctggttgtt
tggcctgctg ccctttca ggggtcctaa gcccacaatc atgcctccct aagaccttgg
catccctccc tctaagccgt tggcacctct gtgccaccc tcacactggc tccagacaca
cagccctgtgc tttggagct gagatcactc gcttcaccct cctcatctt gttctccaag
taaagccacg aggtcggggc gagggcagag gtgatcacct gcgtgtccca tctacagacc
tgcaagctca taaaacttctt gattcttctt cagcttggaa aagggttacc ctggcactg
gcctagagcc tcacccctta atagacitg ccccatgagt ttgctcatgtt gaggcaggact
atttctggca ctgcaagtc ccatgatttc ttccgttaatt ctgagggtgg ggggagggac
atgaaatcat cttagcttag ctccctgtct gtgaatgtct atatagtgtt ttgtgtgtt
taacaaatga ttacactga ctgtgtgtt aaaagtgaat ttggaaataa agttattact
ctgatt
ctgtgtacaa ttaatcatcg gctcgataa tgtgtggaaat tgtgagcggtaacaattc
acacagggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctggg

ggaccgtca gtcttcctct tccccccaaa acccaaggac accctcatga tctccggac
ccctgaggtc acatgcgtgg tggtgacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtagtgc gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagtac aagtgcagg tctccaacaa agccctccc gcccccattcg agaaaaccat
ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc catccggga
ggagatgacc aagaaccagg tcagcgtac ctgcctggc aaaggcttct atccagcga
catgcccgtg gagtgggaga gcaatggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttcttcct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggAACGCTC tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 9

atgcctactt caagttctac aaagaaaaaca cagctacaac tggagcattt actgctggat
ttacagatga tttgaatgg aattaataat tacaagaatc ccaaactcac caggatgctc
acatthaagt ttacatgcc caagaaggcc acagaactga aacatctca gtgtctagaa
gaagaactca aacctctgga ggaagtgcta aatttagctc aaagcaaaaa cttdactta
agacccaggg acttaatcag caatatcaac gtaatagttc tggactaaa gggatctgaa
acaacattca tgtgtgaata tgctgatgag acagcaacca ttgttagaatt tctgaacaga
tggattacct ttctcaaag catcatctca acactgactt gataa ggc ggc ggc ggc
aatgtccoga attcccaagcc tcaccacccc ttctcagtaa tgaccctggg tggttgcagg
aggtacctac tccatactga gggtaaaatt aagggaaaggc aaagtccagg cacaagagt
ggaccccagc ctctcactt cagttccact catccaactg ggaccctcac cacgaatctc
atgatctgat tcggttccct gtctcctcct cccgtcacag atgtgagcca gggcactgct
cagctgtac ccttaggtgtt tctgccttgt tgacatggag agagccctt cccctgagaa
ggcctggccc ttccctgtgc tgagcccaca gcagcaggct ggggtcttg ttgtcagtg
gtggcaccag gatggaaggg caaggcaccc agggcaggcc cacagtcccg ctgtccccca
cttgcaccct agcttgtac tgccaacctc ccagacagcc cagcccgctg ctcagctcca
catgcatagt atcagccctc cacacccgac aaagggaaac acacccctt gaaatggtt
ctttcccccc agtcccagct ggaagccatg ctgtctgtc tgctggagca gctgaacata
~~tacatzagat~~ ttgcctgccc ctccccatct gcaccctgtt gagtttagt tggattgtc
tgtttatgt tggattcacc agagtacta tgatagtgaa aaaaaaaaaaaaaaaa
aggacgcacatg tatcttgaaa tgctgtaaa gaggttcta acccaccctc acgaggtgtc
tctcacccccc acactggac tcgtgtggcc tgggtgtgc caccctgctg gggcctcccc
agtttgaaa ggcttcctc agcacctggg acccaacaga gaccagcttc tagcagctaa
ggaggccgtt cagctgtac gaaggcctga agcacaggat taggactgaa gcgatgtatgt
ccccctccctt acttccctt gggctccct gtgtcaggc acagactagg tcttgcggct
ggtctggctt gcggcgcgag gatggttctc tctggcata gcccgaagtc tcatggcagt
cccaaaggag gcttacaact cctgcacatcac aagaaaaagg aagccactgc cagctgggg

gatctgcagc tcccagaagc tccgtgagcc tcagccaccc ctcagactgg gttcctctcc
aagctcgccc tctggagggg cagcgcagcc tcccaccaag ggccctgcga ccacagcagg
gattggatg aattgcctgt cctggatctg ctctagaggc ccaagctgcc tgcctgagga
aggatgactt gacaagtca gagacactgt tcccaaagcc ttgaccagag cacctcagcc
cgctgaccc gcacaaaactc catctgctgc catgagaaaa gggaaagccgc ctttgcaaaa
cattgctgcc taaagaaaact cagcagccctc aggcccaatt ctgccacttc tggttgggt
acagttaaag gcaaccctga gggacttggc agtagaaatc cagggccctcc cctggggctg
gcagcttcgt gtgcagctag agcttacct gaaaggaagt ctctggccc agaactctcc
accaagagcc tccctgcccgt tcgctgagtc ccagcaattc tcctaagttg aaggatctg
agaaggagaa gaaaaatgtgg ggttagattt gttgtggta gagatatgcc cccctcatta
ctgccaacag ttccggctgc atttcttcac gcacccctgg tcccttcctt gaagttctt
tgccctgctc ttccggccatca tggcccttct tatacggaa gctctggat ctccccctt
tggggcaggc tcttggggcc agcctaagat catggtttag ggtgatcagt gctggcagat
aaattgaaaa ggcacgctgg ctgtgatct taaatgagga caatcccccc agggctggc
actcctcccc tccctcact tctcccacct gcagagccag tgtccttggg tggcttagat
aggatatact gtatgccggc tccctcaagc tgctgactca ctttatcaat agttccattt
aaattgactt cagtggtgag actgtatcct gtttgctatt gcttggtgt ctatgggggg
aggggggagg aatgtgtaag atagttaca tggcaaaagg gagatctgg ggtgcagcac
ttaaactgcc tcgtaaccct tttcatgatt tcaaccacat ttgcttagagg gaggagcag
ccacggagtt agaggccctt ggggttctc tttccactg acaggcttc ccaggcagct
ggctagttca ttccctcccc agccagggtgc aggcttagga atatggacat ctggttgtt
tggcctgctg ccctcttca ggggtcctaa gcccacaatc atgcctccct aagaccttgg
catccttccc tctaagccgt tggcacctt gtgccaccc tcacactggc tccagacaca
cagcctgtgc tttggagct gagatcactc gcttcacccct cctcatctt gtctccaag
taaagccacg aggtcggggc gagggcagag gtgatcacct gcgtgtccca tctacagacc
tgcagctca taaaacttct gatttctt cagcttgaa aagggttacc ctgggcactg
gcctagagcc tcacccctta atagacttag ccccatgagt ttgccatgtt gagcaggact
atttctggca ctggcaagtc ccatgattt ttcggtaatt ctgagggtgg ggggagggac
atgaaatcat cttagcttag cttagtgcgt gtgaatgtct atatagtgtt ttgtgtgtt
taacaaatga ttacactga ctgttgctgt aaaagtgaat ttggaaataa agttattact
ctgatt
ctgtgacaa ttaatcatcg gctcgataa tgtgtgaaat tgtgagcggta taacaatttc
acacagggaaa caggatccga taatgacatg cccaccgtgc ccagcacccctg aactcctgg
gggaccgtca gtcttcctct tccccccaaa acccaaggac accctcatga tctccggac
ccctgagggtc acatgcgtgg tggtgacgt gagccacgaa gacccctgagg tcaagttca
ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtctt caccgtccgt caccaggact ggctgaatgg
caaggagttac aagtgcagg tctccaacaa agccctcccc gccccatcg agaaaaccat
ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc catccggga

ggagatgacc aagaaccagg tcagcctgac ctgcctggc aaaggctct atcccagcga
catcgccgtg gagtggaga gcaatggca gccggagaac aactacaaga ccacgcctcc

cgtgctggac tccgacggct ctttccttct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggAACgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 10

tgctgcgcaa cacgatgaag ctcaacaaaa cgcttttat caagtctaa atatgcctaa
cttaaatgct gatcaacgca atggtttat ccaaagcctt aaagatgatc caagccaaag
tgctaacgtt ttaggtgaag ctaaaaaatt aaacgaatct caagcaccga aagctgacaa
caattcaac aaagaacaac aaaatgcitt ctatgaaatc ttgaacatgc ctaactgaa
cgaagaacaa cgcaatggtt tcatccaaag cttaaaagat gacccaagtc aaagtctaa
cctattgtca gaagctaaaa agttaatga atctcaagca ccgaaagcgg ataacaatt
caacaaagaa caacaaaatg cttctatga aatcttacat ttacctaact taaacgaaga
acaacgcaat ggttcatcc aaagcctaaa agatgaccca agccaaagcg ctaaccttt
agcagaagct aaaaagctaa atgatgcaca agcaccaaaa gctgacaaca aattcaacaa
agaacaacaa aatgcttct atgaaattt acatttacat aacttaactg aagagcaacg
taacggcttc atccaaagcc ttaaagacga tccttcagtg agcaaagaaa ttttagcaga
agctaaaaag ctaaacgtatg ctcaagcacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc aacaaacctg gcaaagaaga cggtaac ggc ggc ggc ggc
gcggccgcgt cgaccgcgn cgcgtcgacg gcaaaggaga ctgctgaggc ttttgctgat
ganatactgg anaaggctgg gccacttggt gtgtgtctg ctgttgact tgatataact
gcctaccctt aaaagccaaa
ctgtgacaa ttaatcatcg gctcgatataa tgtgtggat tgtgacgcga taacaatttc
acacaggaaa cagatccga taatgacatg cccaccgtgc ccagcacctg aactcctgg
gggaccgtca gtctccctt tccccccaaa acccaaggac accctcatga tctccggac
ccctgaggc acatgcgtgg tggtggacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtagtgc gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggaggact
caacagcact taccgggtgg tcagcgtctt caccgtccgt caccaggact ggctgaatgg
caaggagttac aagtgcacgg tctccaacaa agccctccca gccccatcg agaaaaccat
ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc catccggga
ggagatgacc aagaaccagg tcagcctgac ctgcctggc aaaggctct atcccagcga
catcgccgtg gagtggaga gcaatggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ctttccttct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggAACgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 11

tgctgcgcaa cacgatgaag ctcaacaaaa cgcttttat caagtctaa atatgcctaa
cttaaatgct gatcaacgca atggtttat ccaaaggcct aaagatgatc caagccaaag

tgctaacgtt ttaggtgaag ctaaaaaatt aaacgaatct caagcaccga aagctgacaa
caatttcaac aaagaacaac aaaatgctt ctatgaaatc ttgaacatgc ctaactgaa
cgaagaacaa cgcaatggtt tcataccaaag cttaaaagat gacccaagtc aaagtctaa
cctattgtca gaagctaaaa agttaatga atctcaagca cggaaagcgg ataacaatt
caacaaagaa caacaaaatg cttctatga aatcttacat ttacctaact taaacgaaga
acaacgcaat ggttcatcc aaagcctaaa agatgaccca agccaaagcg ctaaccctt
agcagaagct aaaaagctaa atgatgcaca agcacaaaaa gctgacaaca aattcaacaa
agaacaacaa aatgcttct atgaaattt acatttacct aacttaactg aagagcaacg
taacggcttc atccaaagcc taaaagacga tccttcagtg agcaaagaaa ttttagcaga
agctaaaaag ctaaacgatg ctcaagcacc aaaagaggaa gacaacaaca aacctggtaa
agaagacggc aacaaacctg gcaaagaaga cgtaac ggc ggc ggc ggc
gtttaggtca cagtgtgtc gatatcacca aggtggctag aagacatgatc atgtctccct
ttcctctgac atctatggac aaagcctta tcacagtccct ggagatgact ccggtgcttg
ggacagaaat catcaattac cgagatggaa tggggcgagt ctttgctcaa gatgtatatg
caaaagacaa ttacccccc ttcccagcat cagtaaaaga tggctatgct gtccgagctg
ctgatggccc aggagatgt ttcatcattt gggatccca agctggtgaa cagccaaactc
agacagtaat gccaggacaa gtcatgcggg ttacaacagg tgctccaata ccctgcggtg
ctgatgcagt agtacaagtg gaagataccg aacttatcag ggaatcagat gatggactg
aagaactga agtgcattt ctggtcaag ctcggccagg ccaagatatc agacccatcg
gccatgacat taaaagaggg gaatgtttt tggccaaagg aacccacatg ggccctcag
agattggctt tctggcaact ttaggtgtca cagaggttga agttaataag ttccagtgg
ttgcagtcat gtcaacaggg aatgagctgc taaatcctga agatgaccc ttaccaggga
agattcgaga cagcaatcgt tcaactttc tagcaacaat tcaggaacat ggttacccca
cgatcaactt gggtattgtt ggagacaacc cagatgactt actcaatgcc ttgaatgagg
gtatcagtgc tgctgtgtc atcatcacat caggggggtt atccatgggg gaaaaggact
atctcaagca ggtgctggac attgatcttc atgctcagat ccattttggc agggttttta
tggaaaccagg ctggccaaaca acatttgcac cttggatat tgatgggttta agaaaaataa
tctttgcact acctggaaat cctgtatcgg ctgtggtcac ctgcaatctc ttgttgtgc
ctgcactgag gaaaatgcag ggcacatctgg atcctggcc aaccatcatc aaagcaaggt
tatcatgtga tgtaaaactt gatcctcgtc cagaatacca tcgggttata ctaacttggc
atcaccacca accactaccc tgggcacaga gtacaggtaa tcaaatgagc agccgtctga
tgagcatgcg cagtgcacat ggattgttga tgctaccccttcc aaagacagaa cagtagtgg
agctccacaa aggccgggtt gtggatgtca tggcattgg acggctatga tggcaccag
ctgttgacaa ttaatcatcg gctcgatataa tgtgtggaaat tgtgagcggtaaacaattt
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacccctg aactcctggg
gggaccgtca gtcttcctct tccccccaaa acccaaggac accctcatga tctccggac

ccctgaggc acatgcgtgg tggggacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtagtg gacggcggtgg aggtgcataa tgccaagaca aagccgcggg aggagcagta
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctgaatgg
caaggagtac aagtcaagg tctccaacaa agccctccc gccccatcg agaaaaccat

ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc catccggga
ggagatgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct atcccagcga
catgccgtg gagtgggaga gcaatggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct cttcttcct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggAACgtct tctcatgctc cgtgatgcat gaggctctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc

Seq. 12

tgctgcgcaa cacgtgaag ctcaacaaaa cgcttttat caagtctaa atatgcctaa
cttaaatgct gatcaacgca atggtttat ccaaaggctt aaagatgatc caagccaaag
tgctaacgtt ttaggtgaag ctaaaaattt aaacgaatct caagcaccga aagctgacaa
caattcaac aaagaacaac aaaatgctt ctatgaaatc ttgaacatgc ctaactgaa
cgaagaacaa cgcaatggtt tcatccaaag cttaaaagat gacccaagtc aaagtctaa
cctattgtca gaagctaaaa agttaatga atctcaagca ccgaaagcgg ataacaatt
caacaaagaa caacaaaatg cttctatga aatcttacat ttacctaact taaacgaaga
acaacgcaat gtttcatcc aaagcctaaa agatgaccca agccaaagcg ctaaccttt
agcagaagct aaaaagctaa atgatgcaca agcacaaaaa gctgacaaca aattcaacaa
agaacaacaa aatgcttct atgaaatttt acatttacct aacttaactg aagagcaacg
taacggcttc atccaaagcc taaaagacga tcctcagtg agcaaagaaa ttttagcaga
agctaaaaag ctaaacgatg ctcaagcacc aaaagagggaa gacaacaaca aacctggtaa
agaagacggc aacaaacctg gcaaagaaga cggtaac ggc ggc ggc ggc
aatgtcccgaa attcccgacc tcaccacccc ttctcagtaa tgaccctgg tgggtgcagg
aggtacctac tccatactga gggtaaattt aaggaaaggc aaagtccagg cacaagagt
ggaccccagc ctctcactct cagttccact catccaactg ggaccctcac cacgaatctc
atgatctgat tcggttccct gtctctcct cccgtcacag atgtgagcca gggcactgct
cagctgtac cctaggtgtt tctgcctgt tgacatggag agagccctt cccctgagaa
ggcctggccc cttccctgtgc tgagcccaca gcagcaggct ggggtcttg gttgtcagt
gtggcaccag gatggaaggg caaggcaccc agggcaggcc cacagtcccg ctgtccccca
cttgcaccct agcttgtac tgccaacctc ccagacagcc cagccgcgtg ctcagctcca
catgcatagt atcagccctc cacacccgac aaaggaaac acacccctt ggaaatggtt
ctttcccccc agtcccagct ggaagccatg ctgtctgttc tgctggagca gctgaacata
tacatagatg ttgccctgcc ctcccatct gcaccctgtt gagttgtatg tggattgtc

tgttatgct tggattcacc agagtacta tgatagtcaa aagaaaaaaaaaaaaaaa
aggacgcacg tatcttgcggaa tgcttgtaaa gaggttcta acccaccctc acgagggtgc
tctcacccccc acactgggac tcgtgtggcc tgggtggc caccctgctg gggcctccca
agtttgaaa ggcttcctc agcacctggg acccaacaga gaccagctc tagcagctaa
ggaggccgtt cagctgtac gaaggcctga agcacaggat taggactgaa gcgtatgt
ccccctccct acttccctt gggctccct gtgtcaggc acagactagg tcttggc
ggtctggc ttgcggc gagatggctc tctggcata gcccyaagtc tcatggc
cccaaaggag gcttacaact cctgcatcac aagaaaaagg aagccactgc cagctgggg
gatctgcacg tcccagaagc tccgtgagcc tcagccaccc cttagactgg gttccctcc
aagctcgccc tctggagggg cagcgcagcc tcccaccaag gcccctgcga ccacagcagg
gattggatg aattgcctgt cctggatctg ctctagaggc ccaagctgcc tgcctgagga
aggatgactt gacaagtcag gagacactgt tcccaaagcc ttgaccagag cacctcagcc
cgctgacctt gcacaaactc catctgctc catgagaaaa gggaaagccgc ctttgc
cattgctgcc taaagaaaact cagcagcctc aggcccaatt ctgccacttc tggttgggt
acagttaaag gcaaccctga gggacttggc agtagaaatc cagggcctcc cctggggctg
gcagcttgtt gtgcagctag agcttaccc gaaaggaagt ctctggccccc agaactctcc
accaagagcc tccctgcccgt tcgtgagtc ccagcaattc tcctaagttt aaggatctg
agaaggagaa ggaaatgtgg ggttagattt gttgggttta gagatatgcc cccctcatta
ctgccaacag ttccggctgc atttcttac gcacccctgtt tcctttccctt gaagttctt
tgccctgctc ttccggccatca tggcccttct tatacggaaag gctctggat ctccccctt
tggggcagggc tcttggggcc agcctaagat catggtttag ggtgatcagt gctggc
aaattgaaaaa ggcacgctgg ctgtgatct taaatgagga caatcccccc agggctggc
actcctcccc tccctcact tctccaccc gcagagccag tgtccctggg tggcttagat
aggatataact gtatgccggc tcctcaagc tgctgactca cttaatcaat agttccattt
aaattgactt cagtggttag actgtatccct gtttgcattt gcttgcattt ctatgggggg
aggggggagg aatgtgtaaat atgttaaca tggccaaagg gagatctgg ggtgc
ttaaactgcc tcgttaaccct ttcatgatt tcaaccacat ttgccttaggg gagggagc
ccacggagttt agaggccctt ggggtttctc ttccactg acaggcttc ccaggc
ggctgttca ttccctcccc agccagggtgc aggcgttagga atatggacat ctgggt
tggccctgctg ccctttca ggggtctaa gcccacaatc atgcctccct aagacctgg
catcctccc tctaagccgt tggcacctt gttccaccc tcacactggc tccagacaca
cagccgtgc ttggtagt gatcactc gttccaccc cctcaatctt gttctccaag
taaagccacg aggtcggggc gagggcagag gtgatcacct gcgtgtccca tctacagacc
tgcagcttca taaaacttctt gatttctt cagtttgc gaa aagggttacc ctggc
gccttagagcc tcacccctta atagacttag ccccatgagt ttgcctgtt gagcaggact
atttctggca ctgcgttca ccatgatttcc ttcggtaattt ctgagggtgg ggggagg
atgaaatcat cttagctttag ctgtgttctt gtgtatgtt atatgtgtt ttgtgtt
taacaaatga ttacactga ctgtgttca aaaagtgaat ttggaaataa agttattact
ctgatt

ctgttacaa ttaatcatcg gctcgataa tgtgtggaaat tgtgagcgga taacaatttc
acacaggaaa caggatccga taatgacatg cccaccgtgc ccagcacctg aactcctgg
gggaccgtca gtctcccttccccc acccaaggac accctcatga tctcccgac
ccctgaggc acatgcgtgg tggtgacgt gagccacgaa gaccctgagg tcaagttcaa
ctggtaacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg aggagcaga
caacagcacg taccgggtgg tcagcgtcct caccgtcctg caccaggact ggctaatgg
caaggagtac aagtgcagg tctccaacaa agccctccca gcccccattcg agaaaaccat
ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc catccgg
ggagatgacc aagaaccagg tcagcctgac ctgcctggc aaaggcttct atcccagcga
catcgccgtg gagtggaga gcaatggca gccggagaac aactacaaga ccacgcctcc
cgtgctggac tccgacggct ccttcttcct ctatagcaag ctcaccgtgg acaagagcag
gtggcagcag gggAACGTCT tctcatgctc cgtgatgcat gaggtctgc acaaccacta
cacgcagaag agcctctccc tgtccccggg taaataatag gatcc